DOI: 10.1007/s10272-014-0517-8

Manfred Gärtner

Standards Are Poor

On Competence and Professional Integrity at the Leading Rating Agency

Competence behind sovereign ratings is crucial, given that the government bond market may be vulnerable to multiple equilibria and self-fulfilling prophecies. With this in mind, this paper reviews and scrutinises official Standard & Poor's publications that address key issues surrounding the government bond market and the role of ratings. It presents a lack of competence, revealed in an inadequate grasp of crucial concepts such as multiple equilibria and self-fulfilling prophecy, obliviousness to Standard and Poor's own rating methodology, and a nonchalant treatment of facts that casts a poor light on the agency's integrity.

The mistakes of a major credit rating agency (CRA) may have serious consequences for the financial sector, the economy and society at large. This is virtually mandated by financial industry regulations but is also due to (and boosted by) financial markets' reliance on the competence of rating agencies. Mistakes cannot be avoided, as our understanding of a continuously evolving world remains imperfect, and economies and political processes are subject to a constantly changing barrage of shocks that appear impossible or very difficult to anticipate. However, an appropriately designed incentive structure that eliminates conflicts of interest and a high level of expertise seem to constitute necessary conditions for keeping rating errors, and the fallout that these may generate, to a minimum. With this in mind, this paper examines the expertise behind sovereign ratings at the world's leading rating agency.

Assessing the sovereign rating competence of CRAs

When assessing the quality of ratings, it is mandatory to separately consider each area in which CRAs engage in business. Crudely speaking, CRAs evaluate the creditworthiness of businesses, the riskiness of financial products and the solvency of sovereigns.

While the first area has a century-old tradition and is at the very origin of the rating industry, the other two are rather new and require their own idiosyncratic kind of expertise. Gauging the risks of complex derivatives, which have flooded the financial markets in recent decades, is completely different from judging the business prospects of commercial companies or the creditworthiness of a country.

Manfred Gärtner, University of St. Gallen, Switzerland.

This paper looks at the professional competence behind *sovereign credit ratings*. While the rating of sovereigns had already been popular for a spell between the two world wars,[1] it had largely vanished after the Second World War. As Figure 1 shows, the big three CRAs – Fitch Ratings, Moody's Investors Service and Standard & Poor's Ratings Services – rated no more than about a dozen sovereigns, if any, before the collapse of the communist Eastern bloc in the late 1980s. Thereafter, this business segment took off, and now each of the world's biggest CRAs rates more than 100 countries.

There appear to be three basic ways to evaluate the quality of sovereign credit ratings: by evaluating the forecasting quality of sovereign ratings, by gauging the consistency in sovereign ratings over time and among countries, and by appraising the expertise of the professionals who rate sovereigns.

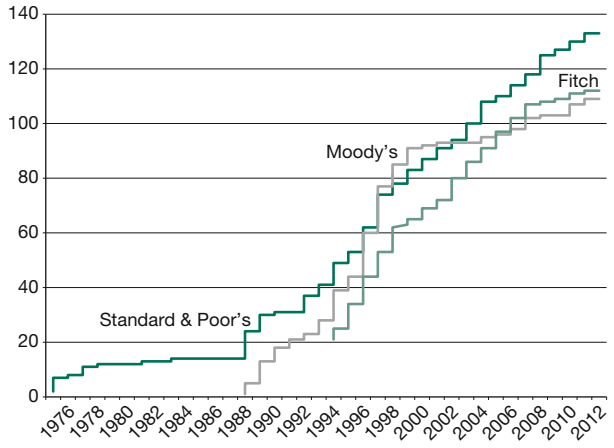
Evaluating the forecasting quality of sovereign ratings

While this appears to be the most natural approach, it does not work – for two reasons. First and foremost, there are no data, or no sufficient data, available that would permit this kind of statistical exercise. Until the Greek haircut in 2012, no industrial country with a credit rating by Fitch, Moody's or Standard & Poor's had defaulted in modern times.[2] This problem was even conceded by ratings agencies well before the financial crisis:

1 See N. Gaillard: A Century of Sovereign Ratings, New York 2012, Springer.

2 While there are a limited number of developing countries with credit ratings that have had to default or restructure their debt in recent years and decades, this number remains insufficient for serious statistical analysis. Additionally, the question of comparability remains; whether studying the defaults of developing countries generates useful lessons for the rating and solvency issues of high income countries that operate in completely different institutional settings is questionable.

Figure 1
The number of countries with a sovereign debt rating by the big three rating agencies



Source: Author's calculations from the web pages of Fitch, Moody's and Standard & Poor's.

> It is important, though, that investors realise the limitations of this [sovereign debt rating] exercise, which is necessarily far less certain than our ability to analyse either bank or corporate risks of default. The essential problem is that the world of sovereign borrowers is far smaller than the world of large banks or corporations, and that the number of instances of default in the modern period when we have reasonable national accounts is tinier still.[3]

Second, even if we had a sufficient number of sovereign defaults, we would still have to deal with the problem of self-fulfilling prophecy. There is a distinct possibility that several equilibria exist in the market for a country's government bonds.[4] Once the perceived default risk crosses a given threshold, initially unfounded surges in risk perception may generate their own justification in terms of increasing interest rates, rising debt or receding growth. The same holds for shocks to the interest rate. In such a situation, even sovereign ratings derived from the roll of a die would be warranted by subsequent events and would appear to be competently generated forecasts.

Gauging the consistency of sovereign ratings procedures over time and among countries

This is a much softer criterion, since consistency reveals nothing about the quality of a procedure. Since the first option is not available, however, this approach has been considered in a number of empirical papers, with uniformly negative results. This line of research has yielded the following insights:

- Sovereign ratings are pro-cyclical and may aggravate financial and debt crises.[5]
- Ratings are biased against debt titles issued by governmental entities.[6] They receive ratings that are about four notches worse than those of private-sector instruments in the same risk category.
- Downgrades of the eurozone's peripheral countries look excessive and cannot be justified by deteriorating fundamentals.[7]
- Sovereign ratings suffer from a cultural and home bias. They favour countries where CRAs are domiciled and English speaking countries in general.[8] These biases have increased during the financial and debt crises.

CRAs have taken issue with this approach and thus avoided the questions posed. Accepting their criticism, if only for argument's sake, would force us to retreat to a still weaker criterion.

Assessing the qualifications and expertise of the people who rate sovereigns

This is a very weak test. Even if the sovereign ratings departments of CRAs were exclusively manned by properly educated geniuses, conflicts of interest would still loom and might lead to rating biases or even abuses. On the other hand, without the qualifications needed to understand the idiosyncrasies of the market for government bonds and how these are related to the economy and the political arena at large, even the most benevolent CRA could not carry out the task properly. Therefore, we are looking at a *necessary condition* for high-quality sovereign debt ratings, albeit not at a sufficient one.

3 See Fitch Ratings: Sovereign Ratings, Rating Methodology, 2002, http://www.fitchratings.com.bo/UpLoad/methodology.pdf, retrieved 31 January 2012, pp. 3 f.

4 D. Romer: Advanced Macroeconomics, McGraw-Hill, 2011.

5 G. Ferri, L.-G. Liu, J.E. Stiglitz: The procyclical role of rating agencies: evidence from the East Asian crisis, in: Economic Notes, Vol. 28, No. 3, 1999, pp. 335-355.

6 J. Cornaggia, K. Rogers Cornaggia, J. Hund: Credit Ratings across Asset Classes: A ≠ A?, SSRN eLibrary, 2012.

7 M. Gärtner, B. Griesbach, F. Jung: PIGS or Lambs? The European Sovereign Debt Crisis and the Role of Rating Agencies, in: International Advances in Economic Research, Vol. 17, No. 3, 2011, pp. 288-299; M. Gärtner, B. Griesbach: Rating Agencies, Self-Fulfilling Prophecy and Multiple Equilibria? An Empirical Model of the European Sovereign Debt Crisis 2009-2011, 2012, Discussion paper No. 2012-15, University of St. Gallen, June 2013, http://www1.vwa.unisg.ch/RePEc/usg/econwp/EWP-1215.pdf; M. Gärtner, B. Griesbach, G. Mennillo: The near-death experience of the Celtic Tiger: A model-driven narrative from the European sovereign debt crisis, in: Intereconomics, Vol. 48, No. 6, 2013, pp. 358-365; D. Vernazza, E.F. Nielsen, V. Gkionakis: The Damaging Bias of Sovereign Ratings, UniCredit Global Themes Series 21, 26 March 2014.

8 A. Fuchs, K. Gehring: The Home Bias in Sovereign Ratings, Discussion Paper Series No. 552, University of Heidelberg, December 2013.

It has already been questioned whether CRAs possess the qualifications needed for competent analysis of sovereign finances. Gaillard noted that "(t)here is a lack of macroeconomic expertise at Fitch, Moody's and Standard & Poor's".[9] More recently, the European Securities and Markets Authority (ESMA) gave a warning and demanded remedial action from CRAs: "ESMA also observed the increasing reliance that lead analysts place on very junior support staff. ESMA is concerned that … this could pose a risk to the quality of ratings.[10]

The current paper takes these concerns seriously. It looks at the issue from a slightly different angle, however, by scrutinising an official contribution by Standard & Poor's (henceforth referred to as S&P) to the ongoing discussion about the role played by the CRAs in the so-called European sovereign debt crisis. This contribution[11] was written and disseminated in reaction to the questions raised in Gärtner and Griesbach (henceforth referred to as GG).[12] A brief summary of GG should prove helpful when evaluating the arguments advanced by S&P.

The GG paper in a nutshell

The GG paper starts with a sketch of Romer's model of the market for government bonds,[13] which conveys the essence of Calvo's optimising model of sovereign debt crises by means of two structural equations.[14] The first equation suggests that rising sovereign default risk raises the interest rate on government debt.

The second equation proposes that rising interest rates, along with negative developments in economic and political fundamentals, increase a country's propensity to default. Depending on the parameters and the state of a country's economy and public finances, this model may generate two equilibria that are locally stable: a good one, in which both the interest rate and default risk are low, and a bad one, in which interest rates rise to unsustainable levels and the government defaults. An insolvency threshold (i.e. a third, *unstable* equilibrium) separates the neighbourhoods within which these two equilibria are stable. Once either the interest rate or the perceived default risk cross their respective threshold values, a self-propelling process is initiated which drives the country closer and closer to insolvency.

Next the paper explores whether this model fits empirical data and may shed light on Europe's sovereign debt crisis. For this purpose, two assumptions are added:

(i) *Financial investors rely on sovereign debt ratings as an indicator of default risk*.

This permits the use of sovereign ratings data as a stand-in for perceived default risk. It also brings CRAs into the picture, which is interesting in itself in the light of concerns about their role during this crisis.[15]

(ii) *Changes in the interest rate may affect default risk with a lag*.

This derives from the observation that most countries' debt titles have long maturities that extend well beyond the one-period horizon proposed, for simplicity, in the Romer model. It also acknowledges that financial markets may not always behave rationally, particularly during times of crisis.[16]

Having set up this augmented model, displayed in Figure 2, the paper attempts to derive the slopes and curvatures of the two relationships from empirical data for 25 OECD countries between 1999 and 2011. The paper reaches four main conclusions.

First, ratings respond linearly to changes in the interest rate. This results in the straight *rating line* shown in Figure 2. As mentioned, interest rates are allowed to affect ratings with a

9 N. Gaillard: Interview, France 24, 2011, http://www.france24.com/en/20110901-interview-norbert-gaillard-economist--world-bank-credit-ratings-agencies-sovereign-debt-moodys-standard-poors-fitch-us-/.

10 ESMA: Credit Rating Agencies. Sovereign ratings investigation. ESMA's assessment of governance, conflicts of interest, resourcing adequacy and confidentiality controls, 2013, http://www.esma.europa.eu/system/files/2013-1780_esma_identifies_deficiencies_in_cras_sovereign_ratings_processes.pdf, p. 14, retrieved 11 June 2014.

11 M. Krämer: S&P's Ratings Are Not "Self-Fulfilling Prophecy", Ratings Direct, Standard & Poor's Ratings Services, August 2012, http://www.standardandpoors.com/spf/upload/Ratings_US/RatingsDirect_Commentary_1003411_08_31_2012_09_23_36.pdf, retrieved 27 March 2013.

12 M. Gärtner, B. Griesbach, op. cit.

13 D. Romer, op. cit.

14 G.A. Calvo: Servicing the public debt: The role of expectations, in: American Economic Review, Vol. 78, No. 4, 1988, pp. 647-61.

15 For academic contributions which scrutinise the role of credit rating agencies during the recent sovereign debt crisis, see A. Fuchs, K. Gehring, op. cit.; M. Gärtner, B. Griesbach, F. Jung, op. cit.; M. Gärtner, B. Griesbach, G. Mennillo, op. cit.; D. Vernazza, E.F. Nielsen, V. Gkionakis, op. cit. and the papers cited therein. There have also been interventions by politicians (see Reuters: Rating agencies warned to watch their step, 2010, http://blogs.reuters.com/global/2010/04/29/rating-agencies-warned-to-watch-their-step/, retrieved 11 June 2014) and by supervisory bodies (see ESMA, op. cit.).

16 On this latter point, see IMF chief economist O. Blanchard: In Review: Four Hard Truths, iMFdirect, 21 December 2011, http://blog-imfdirect.imf.org/2011/12/21/2011-in-review-four-hard-truths/, retrieved 11 June 2014, who included "financial investors are schizophrenic", "post the 2008-09 crisis, the world economy is pregnant with multiple equilibria—self-fulfilling outcomes of pessimism or optimism" and "perception moulds reality" among the four hard truths he learned from the year 2011. The assumption is also in line with results from behavioural economics, suggesting that financial investors may form expectations adaptively. See e.g. E. Haruvy, Y. Lahav, C.N. Noussair: Traders' Expectations in Asset Markets: Experimental Evidence, in: American Economic Review, Vol. 97, No. 5, 2007, pp. 1901-1920.

lag in order to accommodate the reality of long-maturity debt titles. The line moves to the right when the financial or macroeconomic situation deteriorates.

Second, there is pronounced nonlinearity in how ratings affect interest rates. This is shown by the convex *interest rate line* in Figure 2. Successive downgrades trigger bigger and bigger increases in the interest rate. This curve moves upwards when the risk-free rate increases.

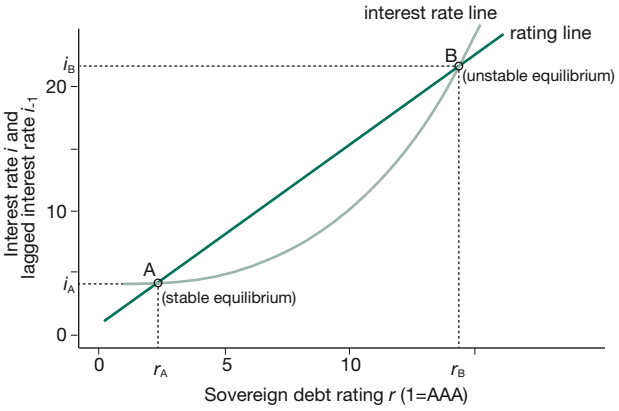
Third, the model may give rise to multiple equilibria and self-fulfilling prophecy. The situation displayed in Figure 2 is synthetic in the sense that sample averages for the exogenous variables position the two lines. There are two equilibria, identified by the points of intersection. In line with the Romer model, the good equilibrium is locally stable. The second equilibrium is unstable, constituting a threshold. Once it is crossed, the interest rate will continually increase, the country's debt will be downgraded to "junk" status and the government will be unable to refinance expiring bond titles, forcing it to default. This situation constitutes a third, stable "equilibrium", even though it is not identified by a point of intersection. On average, for the countries and years included in the data set, the insolvency threshold or abyss appears to lie between the A and the B segment of the rating scale. For individual countries, threshold values may be different, depending on the state of their economy and their public finances.

Finally, several eurozone countries were downgraded "excessively" during 2009-2011, suggesting that the risks posed by the possibility of multiple equilibria and self-fulfilling prophecy in the market for government bonds may be real. Determining whether a downgrade was excessive is accomplished via an econometric equation that links ratings to macroeconomic and financial fundamentals. For example, deteriorating fundamentals would have justified a downgrade of Ireland by one to two rating steps during those two years. In reality, Ireland was downgraded by seven to nine steps by the three leading CRAs.[17]

After a discussion of dynamic extensions of the model, the paper concludes:

> at least for countries with sovereign debt ratings outside the A range even erroneous, arbitrary or abusive rating downgrades may easily generate the very conditions that do actually justify the rating. Combined with earlier evidence that many of the rating downgrades of the eurozone's peripheral countries … could not be justified on the basis of rating algorithms that explain the ratings of other countries or ratings before 2009, this result is highly discomforting. It urges governments to take a long overdue close look at financial markets in general, and at sovereign bond markets in particular, and at the motivations, dependencies and conflicts of interest of key players in these markets.[18]

Figure 2
An empirical model of the market for government bonds



Source: M. Gärtner, B. Griesbach, G. Mennillo: The near-death experience of the Celtic Tiger: A model-driven narrative from the European sovereign debt crisis, in: Intereconomics - Review of European Economic Policy, Vol. 48, No. 6, 2013, pp. 358-365, here p. 359.

The reaction of S&P

Standard & Poor's Rating Services, the world's leading credit rating agency in terms of revenue and employment, published a sweeping defence in response to GG, most succinctly summarised by its heading, *S&P's Ratings Are Not "Self-Fulfilling Prophecies"*.[19] How does S&P substantiate its criticism and denial?

S&P starts by alleging that the GG paper "accuses rating agencies of 'erroneous, arbitrary or abusive rating downgrades' in the context of the European debt crisis". This is as if climate scientists had concluded from a computer-based model that global warming might accelerate if cars did not become more efficient, and Toyota then interpreted this as having been accused of not producing more efficient cars. The passage that S&P registered as an accusation is the implication of a thought experiment, conducted in the context of an empirical model.[20]

Another S&P misperception of what the GG paper expressed and was trying to achieve appears near the end of the S&P critique. There S&P claims that the GG paper "does offer an

17 See M. Gärtner, B. Griesbach, G. Mennillo, op. cit. for a more detailed account of this episode.

18 M. Gärtner, B. Griesbach, op. cit., p. 28.

19 See M. Krämer, op. cit. The critique is signed by Moritz Krämer, Managing Director of Sovereign Ratings for Europe, the Middle East and Africa, which makes it fair to read it as S&P's position. A shorter version of this critique was published in a German economic weekly (M. Krämer: Ratings sind keine „sich selbst erfüllenden Prophezeiungen", in: Wirtschaftswoche, 28 September 2012).

20 M. Gärtner, B. Griesbach, op. cit.

alternative rating approach with its quantitative algorithm" and generously "welcome(s) the competition of ideas and diverse methodologies to assess sovereign credit risk. Whether users of ratings will accept the algorithm remains to be seen, however."

Contrary to S&P's perception, the rating equation presented by GG does not propose an alternative or purportedly superior rating methodology. It does not tell CRAs how they *should* conduct their rating of sovereigns. It offers an interpretation of how they *did* conduct their rating. Employing established methods from empirical economic research, it attempts to retrace what appears to drive actual ratings. Looking through the lens of mainstream macroeconomics, it explores whether and how those variables that are considered crucial for a country's solvency – e.g. the debt ratio, the deficit ratio, economic growth, etc. – affect a country's rating. The resulting equation quantifies how those key variables made their way through the rating procedures of CRAs and provides a description of a core process behind the pertinent work of CRAs, a core process that explains a major part of the rating gaps among countries and of rating changes over time.

The sobering conclusion here is that S&P unfortunately fails to grasp the direction of GG and, hence, misses the questions posed by the paper's results.

Standard & Poor's points of criticism

Expanding on the claim made in its comment's title, S&P maintains that GG's "confusion of correlation and causality is the flawed foundation of their claim that rating actions are self-fulfilling prophecies".

Let us structure this argument. Self-fulfilling prophecies would be triggered and propelled by what S&P calls a "negative-feedback loop".[21] The two segments of such a loop, the existence of which S&P denies, would be the interest rate and the rating equations estimated by GG. These imply the following two links:

Link (1): Deteriorating credit ratings lead to higher interest rates.

Link (2): Rising interest rates have a negative effect on a country's credit rating.

The discomforting property of such an interactive process is that any unjustified increase in the interest rate would trigger a downgrade, which would provide *ex post* justification for the initially unjustified interest rate hike. By the same token, an unjustified downgrade drives up the interest rate, which would provide *ex post* justification for the initially erroneous or arbitrary downgrade.

If quantitative responses are small, this process is *stable* (or self-correcting), meaning that after a shock to the interest rate or to the rating, the market returns to the initial equilibrium. An element of self-fulfilling prophecy remains, but this is not large enough to provide full justification for the initial error (or shock). The effect vanishes over time.

Large responses, however, render the system *unstable* (or self-reinforcing). Any errors committed by CRAs, or by the market in general, become self-fulfilling in full guise, triggering an avalanche of downgrades and interest rate increases that ultimately ends in default.

S&P denies the possibility of self-fulfilling prophecy and, hence, the existence of, in its terminology, a negative-feedback loop in which ratings may "mould reality". This is tantamount to denying link (1) or link (2), or both. Since the statistical correlation between the interest rate and the rating is undisputed, this boils down to the issue of causality. What can we say about links (1) and (2)?

Link (1): Do sovereign ratings affect the interest rate?

In search for an answer, we may consider statistical evidence, take the effects of financial market regulations into account, learn from real-world experiments, and discuss asymmetries, reputation spill-overs and insider information.

Statistical evidence. While the issue of causality between interest rates and sovereign ratings is a tricky matter indeed, there are statistical methods that offer help. One example is the test for Granger causality. Pertinent tests reported by Gärtner, Griesbach and Jung "can never reject the hypothesis that the credit spread [the key driver of which is the interest rate] is caused by the rating".[22]

Of course, such tests have well-known limitations. They never give final proof, one way or the other, but only serve as pieces of evidence. That said, there is a difference between abstract statistical reasoning and econometrics. Econometrics combines the information that we crunch from the numbers with information about the institutions within which the data were generated, with an understanding of the incentive structures and more. Hence, let us look at the issue of causality from a different angle.

21 This terminology is unconventional, because negative feedback is usually associated with self-correcting processes. Self-reinforcing processes, which is what self-fulfilling prophecies amount to, are the result of positive feedback.

22 M. Gärtner, B. Griesbach, F. Jung, op. cit., p. 295.

Regulations. Financial market regulations such as the Basel II accord oftentimes force financial institutions to respond to sovereign downgrades, regardless of whether these institutions consider them justified or not, by adjusting the composition of their portfolios.

Real-world experiments. On occasion, natural experiments teach us more than theoretical and statistical discourses. One such experiment played out on 10 November 2011:

> Standard & Poor's mistakenly announced the downgrade of France's top credit rating on Thursday ... The erroneous alert, which S&P said was sent to some of its subscribers, … contributed to the worst day for France's government bonds since before the euro was launched in 1999. … In a statement issued nearly two hours after the fact, S&P said the message resulted from a technical error and not from any action it intended to take against France.[23]

> As word spread Thursday about the message, the euro weakened against the dollar, U.S. stocks slipped and French bond prices fell, pushing yields higher. Cash reversed course, flooding into U.S. Treasurys and German government bonds.[24]

Asymmetries, reputation transfers and insider information. There are a host of other channels through which sovereign ratings may make their way into the government bonds market:

Asymmetries in the incentive structure of account managers*:* Suppose that Portugal suffers a downgrade that my account manager considers unjustified. While she could recommend that I hold on to my Portuguese government bonds, if Portugal were to default, she would find it hard to justify this advice to me and her boss. If she instead advises me to sell these titles and Portugal does not default, she has the convenient excuse that she only heeded the "opinion" of some well-reputed rating agency.

Actual or imagined insider knowledge: Many sovereign ratings are unsolicited, which means that they are based on publicly available information. Some are solicited, however. And since investors oftentimes cannot or do not discriminate between these two categories, they suspect that some private information was incorporated into any sovereign rating.

Reputation spill-overs: Ratings of companies, in existence for a century, draw on information that the typical investor does not have. This is how rating agencies built their reputations. Many market participants transfer this reputation to the new sovereign rating activities, even though these require an entirely different kind of expertise.

Rationality: When faced with an obviously erroneous downgrade, even the most rational investor will nonetheless react if he believes that other market participants will also respond to the downgrade.

These observations and arguments make it difficult to doubt that sovereign ratings have a causal effect on the interest rate. Let us now turn to the second segment of the feedback loop.

Link (2): Do interest rates affect sovereign ratings?

This is the segment that S&P questions most vehemently:

> Standard & Poor's sovereign methodology does not take credit spreads into consideration. We believe that spread movements provide neither diagnostic capabilities nor do they systematically signal insights into the fundamental factors that differentiate one credit from another.[25]

This is a strong statement. It implies that it is irrelevant whether Japan pays an interest rate of two per cent or 20 percent on its public debt, which is greater than 200 per cent of its GDP. A look at Standard & Poor's sovereign rating methodology gives both comfort and a scare, however:

> Among the "five key factors underlying our sovereign rating analysis" is the "fiscal score" (p. 4). One of two factors determining the fiscal score is a country's "debt burden" (p. 23). And the debt burden is determined by the "General government's interest expenditures as a percentage of government revenues" (p. 27).[26]

The good news is that S&P definitely takes credit spreads (the difference between a country's interest rate and the risk-free rate) into account. The bad news is that S&P is not aware of this. For all its disturbing implications, this borders on the comical, given that S&P "wonder(s) whether the authors [i.e. GG] have even read the [CRAs'] methodologies".[27]

Ignorance of its own rating methodology and a hasty brushing aside of uncomfortable questions plagues other S&P contributions to the ongoing discussion of sovereign debt ratings as

23 Reuters: France shocked by S&P downgrade error, 2011, http://uk.reuters.com/article/2011/11/11/uk-france-ratings-sandp-error-idUKTRE7AA12820111111, retrieved 28 March 2014.

24 Wall Street Journal: S&P 'Oops' on Rating of France Is Probed, 2011, http://online.wsj.com/news/articles/SB10001424052970204224604577030083804142906, retrieved 28 March 2014.

25 M. Krämer: S&P's Rating Are Not … , op. cit. p. 3.

26 Standard & Poor's: Sovereign Government Rating Methodology And Assumptions, 30 June 2011.

27 M. Krämer: S&P's Rating Are Not … , op. cit., p. 3.

well. One recent example is the response to a paper by a team of UniCredit economists.[28] This paper reports unexplained downgrades of countries from the eurozone's periphery similar to those given by GG. Taking issue with UniCredit's claim that "there is little dispute about which variables should be included in an analysis of a country's creditworthiness",[29] S&P's untitled response retorts:

> We dispute UniCredit's analysis and selection of variables. In fact, *none of the ten variables chosen by them is referenced in our own methodology* [emphasis added].[30]

Table 1 lists the ten variables used in UniCredit's analysis on the left. The second column shows pertinent quotes from S&P's methodology. Comparing the two columns, the verdict is clear: The only variable chosen by UniCredit that has no counterpart in the S&P rating manual is nominal GDP. All other nine variables feature in the S&P methodology - either directly, with minor differences in the fine print, or are at least mentioned and discussed as adjustment factors. One example is GDP per capita, which both sources measure in US dollars. UniCredit adjusts for purchasing power. S&P makes adjustment for currency over- and undervaluation, which is essentially the same thing. Another example is public debt. UniCredit uses gross debt while S&P looks at net debt. A final example is law. UniCredit data are from the World Bank Rule of Law index. S&P looks at the "rule of law".

One comes to the disturbing conclusion, then, that S&P's statement is grossly misleading, a fabrication. Apparently, one of S&P's key weapons in defending its sovereign ratings against academic and professional criticism is to mislead readers and the public.

This brings us to S&P's factual claims. As discussed above, a key concern of S&P's critique of the GG paper is the possibility of a feedback loop between the interest rate and the debt rating. Justifying its denial of such a loop and thus ruling out self-fulfilling prophecy, S&P claims that "the [GG] paper offers no support for its claim of a negative-feedback loop".[31] But GG distils exactly such a feedback loop from the data. One may take issue with the result reported in GG, one may call for different empirical methods, insist on the inclusion of other variables, question the time range, the selection of countries, and more. But when S&P asserts that "the authors' claims remain unsupported by evidence",[32] this casts a strange light on their professional integrity.

S&P also maintains:

> In our view, the algorithm [reported in GG] ... would have entirely missed the Greek default in early 2012, the largest sovereign restructuring in financial history. By contrast, far from having acted in an "arbitrary or abusive" manner, Standard & Poor's anticipated Greece's default well before it occurred.[33]

This only reconfirms S&P's misunderstanding of the question being posed. When GG and others ask whether Greece defaulted *because of* downgrades that appeared excessive and which triggered a run of self-fulfilling prophecy, how can the response be to crow that one "predicted" Greece's 2012 default?

Moreover, apart from the possibility of self-fulfilling prophecy, S&P's understanding of anticipating a default "well before it occurred" is quite unusual. Actually, S&P had bestowed upon Greece an A rating a few months after the country joined the eurozone on 1 January 2001. Greece never slipped below this mark until 14 January 2009, when it was downgraded one notch to A-. This happened four months after the collapse of Lehman Brothers, when it had become common knowledge that the financial crisis would wreak havoc with government budgets. This downgrade kept Greece's debt in the investment grade range, which is characterised by the "strong capacity to meet financial commitments, but somewhat susceptible to adverse economic conditions and changes in circumstances".[34] To cite this late and minimal initial downgrade as evidence of S&P's farsightedness is embarrassing.

In the same vein, in its response to Vernazza, Nielsen and Gkionakis,[35] S&P boasts that "S&P was the first rating agency to raise concerns about brewing euro area problems when we started to lower peripheral sovereign ratings, beginning with Italy in mid-2004."[36] In actuality, S&P issued four downgrades and two upgrades of eurozone countries between 2004 and 2008, which hardly classifies as a wake-up call for a currency zone with 15 members (as of 2008). It took until 2009, well after the collapse of Lehman Brothers, when the crisis truly exploded and one crisis summit chased another, for S&P and the other CRAs to start a virtual orgy of procyclical downgrades, during which the downgrades for the eurozone victims massively exceeded what was called for by deteriorating fundamentals. This culminated in 2011 and 2012, with 44 and 43 downgrades

28 D. Vernazza, E.F. Nielsen, V. Gkionakis, op. cit.
29 Ibid., p. 7.
30 Standard & Poor's: Untitled manuscript, 2014, http://ftalphaville.ft.com/files/2014/03/SP-UniCredit-response-27-Mar-14-1.docx, retrieved 29 May 2014.
31 M. Krämer: S&P's Rating Are Not ... , op. cit., p. 2.
32 Ibid.
33 Ibid., p. 4.
34 Standard & Poor's: Credit Rating definitions & FAQs, 2014, http://www.standardandpoors.com/ratings/definitions-and-faqs/en/us, retrieved 6 June 2014.
35 D. Vernazza, E.F. Nielsen, V. Gkionakis, op. cit.
36 Standard & Poor's: Untitled manuscript ... , op. cit.

Table 1
Comparing regressors used in UniCredit paper with S&P methodology

Variables used by UniCredit	Standard & Poor's methodology
Nominal GDP GDP in current prices and exchange rates, USD	*Not mentioned*
GDP per capita GDP per person, PPP-adjusted USD	GDP per capita In USD Standard & Poor's most prominent measure of income levels (p. 16) *Adjustment for over- and undervaluation*
GDP growth Average annual real GDP growth	Real per capita GDP trend growth (p. 17)
Public debt General government gross debt, end-of-year, % of GDP	Debt level Net general government debt as percentage of GDP (p. 27)
Current account Annual current account balance, % of GDP	Current account on average over the last historical year, the current year, and the next two forecast years (p. 20)
External debt Gross external debt, % of GDP	External indebtedness ratio of "narrow net external debt" to current account receipts (p. 19)
Past default Indicator variable that takes the value one in all years following a default event since 1960; zero otherwise	*No direct mention, but indirect references* (A) government's debt payment culture represents a credit risk. (p. 14) (W)ith each successive default, serial defaulters have less of a reputation to lose. (p. 14)
Advanced country An indicator variable that takes the value one if the country is deemed advanced by the IMF; zero otherwise	*No direct mention* *Many factors in discussed in methodology reflect economic and institutional development.* Strength and stability of a country's institutions (p. 7)
Government World Bank Government Effectiveness index	Political score The World Bank's "Worldwide Governance Indicators," which measure six broad dimensions of governance … including … governance effectiveness … (pp. 12-13) Governance and political risk are among the main drivers … that lead to default. (p. 9) strength and stability of the government's institutions, and the effectiveness of its policy-making (p. 7)
Law World Bank Rule of Law index	Rule of law (p. 14)

Claim: "(N)one of the ten variables chosen by [UniCredit] is referenced in our own methodology", Standard & Poor's: Untitled manuscript, 2014, http://ftalphaville.ft.com/files/2014/03/SP-UniCredit-response-27-Mar-14-1.docx, retrieved 29 May 2014.

Note: All terms and text strings in regular typeface are quotes from D. Vernazza, E.F. Nielsen, V. Gkionakis: The Damaging Bias of Sovereign Ratings, UniCredit Global Themes Series 21, 26 March 2014 (left column) or Standard & Poor's: Sovereign Government Rating Methodology And Assumptions, 30 June 2011 (right column). Author's own comments are given in italics.

by the big three CRAs, respectively.[37] During those years, it felt like not a day passed without another downgrade, outlook deterioration, negative watch issue or other bad news about public finances hitting the media.[38] Again, S&P is adorning itself with unearned laurels, making deceptive factual claims.

Conclusion

Competence behind sovereign ratings is crucial, given that the market for government bonds may be vulnerable to multiple equilibria and self-fulfilling prophecies. Several authors have questioned the level of expertise that drives sovereign rating decisions, pointing to a lack of advanced formal education in macroeconomics in the sovereign ratings departments of CRAs[39] and to a tendency to shoulder junior staff with crucial tasks and decisions they may not be up to (yet).[40] The current paper augments earlier criticism by taking a close look at official S&P publications that address issues surrounding the market for government bonds and the role of sovereign ratings. The sobering result is that these contributions display an inability to engage in logical discourse, a lack of understanding of crucial concepts such as multiple equilibria and self-fulfilling prophecy, obliviousness of S&P's own rating methodology, and a nonchalance in making factual claims that casts poor light on the integrity and credibility of S&P. This should be worrisome for anybody whose well-being and future depend on sovereign rating verdicts – which includes virtually all of us.

37 With the big three rating agencies very much marching in lockstep, these downgrades were shared more or less equally among Fitch, Moody's and S&P.

38 The shift looks even more striking when we look at the magnitude of the rating changes published by S&P. Between 2004 and 2008, eurozone member countries received net downgrades – the difference between the sum of the notches (or steps) by which countries were downgraded and the sum of the notches by which they were upgraded – of four notches. By comparison, between 2009 and 2013, eurozone members were dealt net downgrades totalling 59 notches!

39 N. Gaillard, op. cit.

40 ESMA, op. cit.